Volume 328, number 1,2, 169–173 FEBS 12821 August 1993

Purification of the β product encoded by the *Streptococcus pyogenes* plasmid pSM19035

A putative DNA recombinase required to resolve plasmid oligomers

Fernando Rojo[a,b], Frank Weise[c] and Juan C. Alonso[a,b,c]

[a]*Centro de Biología Molecular 'Severo Ochoa', UAM-CSIC, Canto Blanco, 28049-Madrid, Spain.* [b]*Centro Nacional de Biotecnología, CSIC, Canto Blanco, 28049-Madrid, Spain* and [c]*Max-Planck Institut für Molekulare Genetik, Ihnestrasse 73, D-1000 Berlin 33, Germany*

Received 24 May 1993; revised version received 21 June 1993

Genetic evidence suggests that the gene β product of *Streptococcus pyogenes* plasmid pSM19035 is required for converting plasmid multimers into monomers. The β protein was purified from cells overexpressing the cloned gene. N-terminal protein sequence analysis demonstrated that the purified protein had the predicted sequence, except that the N-terminal initiator methionine was not present. Native β protein consists of a dimer of two identical subunits with a molecular mass of 23.8 kDa (25 kDa in SDS-PAGE). The β protein (isoelectric point of 9.7) binds specifically to a DNA fragment (312 bp in length) which contains the promoter region of the orf α-gene β operon and two regions (sites I and II) that show dyad axes of symmetry. It is proposed that protein β binds to sites I and II to mediate resolution of plasmid oligomers.

β Protein (*Streptococcus pyogenes*); DNA binding site; Site specific recombination

1. INTRODUCTION

The Gram-positive broad-host-range *Streptococcus pyogenes* plasmid pSM19035 (27 kb in length) has extraordinarily long inverted repeated sequences that comprise about 80% of the plasmid molecule [1]. The sequences required to ensure duplication of replicons, ordered partition at cell division and DNA inversion, are localized within the inverted repeated segments [2,3]. Hence, these elements are present twice in the plasmid.

Genetic evidence suggests that the β protein might secure pSM19035 partition by catalysing the conversion of dimers or higher order oligomeric forms into monomers, and that it might be also involved in DNA inversion because both events are blocked when the product encoded by gene β is inactivated (see [2]).

Unlike the small high copy plasmids from Gram positive bacteria, pSM19035 uses a unidirectional theta replication mode [3]. A replication origin is located within each arm of the inverted repeated DNA segment [1–3]. If replication proceeds unidirectionally from both replication origins it is expected that it should terminate when the replication forks meet [4]. Under these conditions only a part of the genome is expected to replicate. Such a paradox could be explained by Futcher's model for 2 μm plasmid replication [5] if it is assumed that the plasmid encodes a DNA invertase protein. Indeed, one of the plasmid encoded products, protein β, shows homology with enzymes of the Tn*3*/Hin family of DNA-resolvases and DNA-invertases [2,6,7]. Hence, it is expected that the biochemical characterization of the β product might broaden our insight in the DNA replication mode used by plasmid pSM19035. Furthermore, DNA resolvases bind to a DNA segment containing three binding sites (termed *res*I to III). Each of the *res* subsites consists of two 12-bp long inversely oriented sequences (half site) that are separated by a short space of variable length. Recombination occurs at the centre of *res*I and all three sites are required for efficient recombination (reviewed in [7]). DNA invertases, however, bind to a DNA region that contains a 26-bp segment. The binding site consists of two 12-bp long inversely oriented binding sites (half site) separated by a 2-bp spacer. Here recombination requires an additional *cis*-acting DNA segment (termed enhancer) and the host encoded protein FIS (reviewed in [6]). The biochemical characterization of protein β will help our understanding of site-specific recombination.

In this paper we describe the purification of the β protein from cells overexpressing the gene, the characterization of its physical properties and its interaction with DNA. The results presented provide the first evidence that a protein involved in DNA resolution and

Correspondence address: J.C. Alonso, Centro Nacional de Biotecnología, CSIC, Canto Blanco, 28049-Madrid, Spain. Fax (34) (1) 585 45 46.

DNA inversion binds to a DNA segment containing two putative binding sites (sites I and II).

2. MATERIALS AND METHODS

2.1. *Enzymes and reagents*

Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from Sigma Co. (USA) Ampicillin was purchased from Bayer (Germany). Cellulose phosphate was purchased from Whatman (USA). DNA restriction and modification enzymes, rifampicin and Poly[d(I-C)] were purchased from Boehringer Mannheim (Germany) Ultra pure acrylamide, bis-acrylamide and chloramphenicol were from Serva (Germany). Guanidine HCl was obtained from Schwarz/Mann Biotech ICN (USA) The molecular weight of protein standards was obtained from GIBCO-BRL (USA).

All chemicals used were reagent grade and solutions were prepared in quartz-distilled H_2O

Covalently closed circular plasmid DNA was purified using the SDS lysis method [8] followed by purification on a cesium chloride-ethidium bromide gradient. Labeling of purified DNA subfragments from plasmid pBT233 [2] was performed by filling in the restriction site with the large fragment of the DNA PolI in the presence of [α-^{32}P]dATP and dTTP, dCTP and dGTP as described [8].

The expression of the β protein and the specific labeling with [^{35}S]methionine (Amersham Buchler GmbH, Germany) was carried out as described [2]. The proteins were separated by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) as previously described [9]

2.2. *β Protein purification*

The β protein was purified from *E. coli* BL21(DE3) [10] cells that overproduced it from plasmid pBT241 [2] by monitoring the radiolabeled protein β. Plasmid pBT241 contains gene β downstream of a T7 promoter. *E. coli* cells were lysed and the overexpressed β protein (Fig. 1, lane 2) was readily sedimented by low speed centrifugation (27,000 × *g* for 30 min in buffer A (50 mM Tris-HCl, pH 7.5, 5% v/v glycerol) containing 1 M NaCl).

The pellet was resuspended in buffer A containing 500 mM NaCl (Fraction 1) and centrifuged at 12,000 × *g* for 30 min. The supernatant (Fig. 1, lane 3) was discarded. The pellet, which contains the β protein, was collected and resuspended in buffer A containing 1 M guanidine HCl. The solution was stirred for 15 min at 4°C and centrifuged at 27,000 × *g* for 30 min The supernatant was retained (Fraction 2). The solubilized β protein was refolded under conditions which minimized the formation of protein co-aggregates (low protein concentration (3 $\mu g \cdot ml^{-1}$) and 4°C). A denaturation/renaturation procedure is part of the standard protocol to purify Tn*3* resolvase and does not affect the activity of the protein (see [11]).

The refolded β protein (Fig. 1, lane 4) was loaded onto a phosphocellulose column equilibrated with buffer A containing 250 mM NaCl. The β protein was eluted with a 10-column volume linear gradient from 300 to 500 mM NaCl. The fractions containing the β protein were pooled and applied onto a second phosphocellulose column for concentration. The protein was eluted at 600 mM NaCl in buffer A (Fig. 1, lane 5). The purified β protein, which coincides with the radioactive material (data not shown), was diluted with glycerol (50% v/v final concentration) and stored at −20°C. The purified polypeptide aggregates and precipitates in low-ionic-strength buffers (< 100 mM NaCl) in the absence of DNA at near neutral pH, a behaviour similar to that of the Tn*3* resolvase [7,11].

2.3. *Filter binding assay*

The formation of β protein–DNA complexes was measured by retention on alkali-treated filters (Millipore, type HAWP 0.45 μm) as previously described [12]. Standard reaction (25 μl) was carried out in a solution containing 25 ng of ^{32}P-labeled DNA, 500 ng of Poly[d(I-C)] (as non-specific competitor) and 45 ng (76 nM) of β protein in buffer B (50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$) and incubated during 10 min at 20°C. Ice-cold buffer B (1 ml) was added to stop the reaction, and then the mixture was filtered through KOH-treated filters The filters were dried and the amount of radioactivity bound to the filter was determined by scintillation counting. The value obtained for the DNA retained was corrected for retention in the absence of β protein. The specific activity of the labeled DNA was measured as TCA precipitable material. All reactions were performed in duplicate.

2.4. *Molecular weight determination*

The native molecular weight of β protein was determined by gel filtration FPLC using a Superose 12 column (Pharmacia, Sweden). Chromatography was carried out in buffer A containing 1 M NaCl at 4°C with a flow rate of 0.5 ml/min, and the A_{280} was measured 20 μg of β protein was applied onto the column. A standard curve of K_{av} versus $\log_{10}$ of relative mobility was determined as recommended by Pharmacia. Protein standards were obtained from Pharmacia (chymotrypsinogen A, 25 kDa; BSA, 68 kDa; aldolase, 158 kDa; catalase, 232 kDa, ferritin, 440 kDa; thyroglobulin, 669 kDa and blue dextran, 2,000 kDa)

2.5. *Other methods*

The pI of the β protein was determined using an isoelectric focusing kit as recommended by the suppliers (Servalyt PreNets, Serva, Germany). The N-terminal amino acid sequence was determined by Volker Kruft (Max Planck Institut für Molekulare Genetik, Abt. Wittmann, Berlin, Germany) with an automated Edman degradation in a pulse liquid phase sequencer (model 477A, Applied Biosystems, USA).

The concentration of DNA was determined using the molar extinction coeficient of 6,500 M^{-1} cm^{-1} at 260 mm The β protein concentration was determined using the molar extintion coeficient of 7,740 M^{-1} cm^{-1} at 280 nm

3. RESULTS AND DISCUSSION

3.1. *Purification and physical properties of the pSM19035 β protein*

The β protein was overexpressed using the T7 RNA polymerase expression system [13]. Plasmid pBT241, which contains gene β downstream of a T7 promoter, was introduced into *E. coli* strain BL21(DE3)/pLysS, and gene β was overexpressed by the addition of IPTG and rifampicin as previously described [2]. The β polypeptide, under the expression conditions described in Section 2, accounts for about 8% of total protein mass. Under these experimental conditions a yield of about 2 mg of β protein per litre of induced culture was obtained.

The overproduced 25 kDa polypeptide (predicted molecular mass 23.8 kDa) was insoluble. The β protein aggregates could, however, be dissolved in the presence of 1 M guanidine HCl. This property was exploited in our purification scheme to release unwanted proteins and the remaining chromosomal DNA. After refolding, the β polypeptide was purified by conventional column chromatography monitoring the radiolabeled β protein. Fig. 1 shows the progressive purification of the β polypeptide. The β polypeptide obtained was more than 98% pure, as judged by SDS-PAGE (Fig. 1, lane 5) and quantitative analysis of the N-terminal amino acid (data not shown).

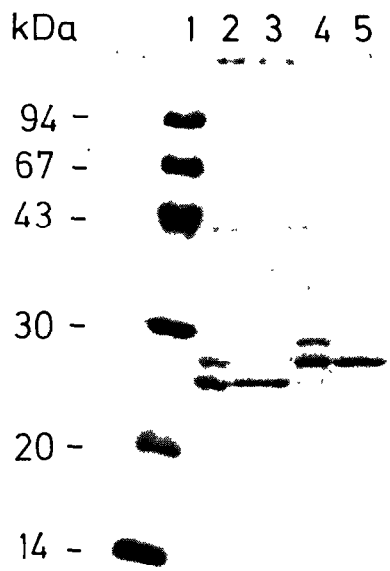

Fig. 1. Purification of β protein. Coomassie blue-stained 15% SDS-PAGE. Purification of protein β is described in the text. Lane 1, molecular weight standards (in kDa); lane 2, cell lysate (fraction 1); lane 3, supernatant of 27,000 × g; lane 4, supernatant after guanidine-HCl treatment (fraction 2); lane 5, phosphocellulose of 460 mM elution (fraction 3).

Identification of the final product as β protein was confirmed by sequencing the amino terminus of the purified protein. The sequence of the first 19 amino-terminal residues of the purified protein is in perfect agreement with the prediction from the nucleotide sequence of the β gene [2], except for the Met residue of the initiator codon which is absent in β specified by pBT241.

The β polypeptide has a high content of charged residues resulting in a calculated isoelectric point of 10.7 [14]. The experimentally determined isoelectric point of the β protein is 9.7 (see Section 2).

The β polypeptide consists of 205 amino acid residues corresponding to a molecular mass of 23,841 Da deduced from the nucleotide sequence, which is slightly less than the one observed after electrophoresis under denaturing conditions and staining with Coomassie blue (M_r 25,000). The molecular weight of native β protein was determined by gel filtration (Fig. 2). The analysis of gel filtration chromatography was carried out by FPLC using a Superose 12 column. From the elution profile of protein β and of a number of protein standards we estimate that the M_r of protein β is aproximately 55,000, twice that of a β gene protomer. Therefore, protein β is probaly a dimer in solution.

3.2. *β Protein binds specifically to DNA*

Genetic evidence suggests that the information necessary for resolving dimers or higher order oligomers of pSM19035 into monomers laid within the 3,092 bp *Hin*dIII DNA fragments B and B′ [1,2]. These DNA segments, which are present twice in the plasmid (within the inverted repeated region, see Section 1), code for the 3′-end of the initiation replication gene (repS), the plasmid replication origin, the orf α-gene β operon and the 5′-end of orf γ in a 5′ to 3′ polarity (Fig. 3) [2,3].

To investigate whether the β protein interacts specifically with DNA, we applied filter binding techniques. The DNA substrate used was the 3,092 bp *Hin*dIII DNA fragment (Fig. 3). *Sty*I digestion of this *Hin*dIII DNA segment generates subfragments of 1,408, 932 and 752 bp in length. The DNA subfragments were purified and labeled by filling-in the ends with [α-^{32}P]dATP. Non-incorporated nucleotides were removed by gel filtration.

^{32}P-labeled DNA was incubated with the β polypeptide in the presence of a 20-fold weight excess of Poly[d(I-C)], and the mixture was passed through a nitrocellulose filter. The β protein was able to mediate the binding of the 1,408 bp DNA subfragment to the nitrocellulose membrane (Fig. 3). The 1,408 bp *Hin*dIII–*Sty*I DNA subfragment was further dissected by digesting the segment with *Ase*I. The β protein was able to form a complex with the purified 815 bp *Ase*I–*Sty*I fragment that is retained on nitrocellulose membrane filters. This

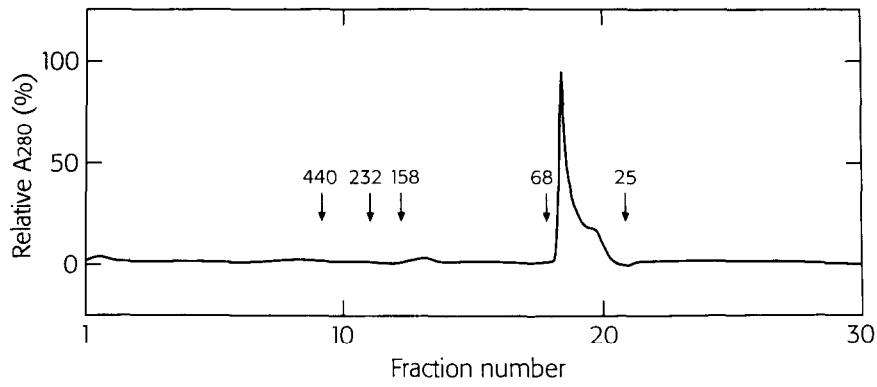

Fig. 2. Molecular weight determination of the purified β protein. Gel filtration chromatography was carried out using a Superose 12 column on a Pharmacia FPLC apparatus. The molecular weight (in kDa) of protein standards is indicated by vertical arrows.

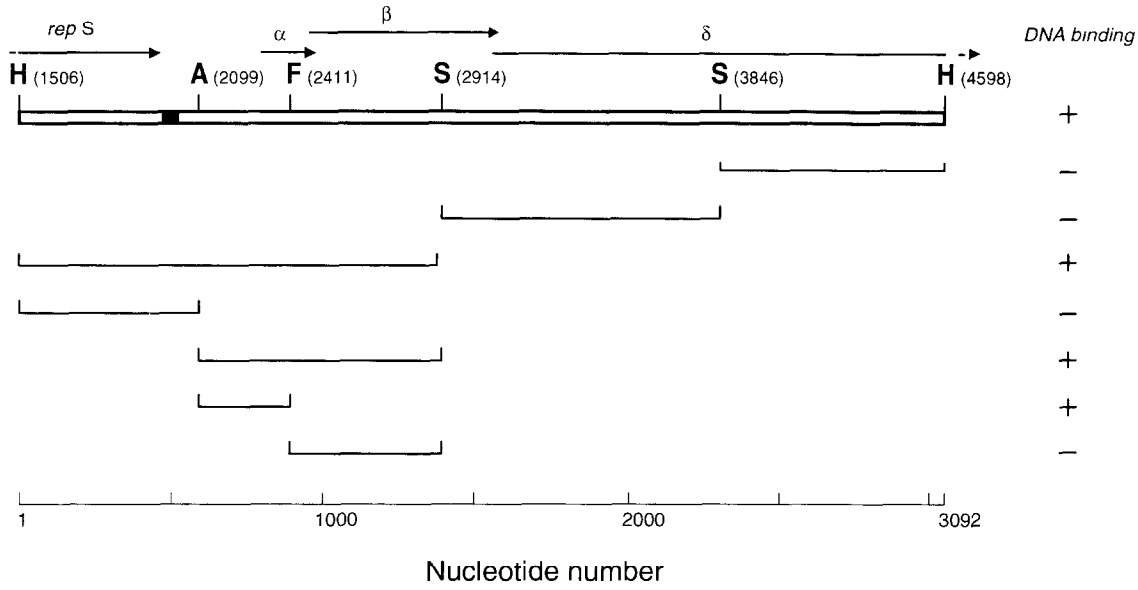

Fig. 3. Analysis of β protein binding to DNA by nitrocellulose filter binding assays. The open bar denotes the physical map of the 3.0 kb *Hin*dIII DNA fragment and the filled box denotes the plasmid replication origin. The locations of the genes are denoted by arrows (see text). Thin lines denote the purified DNA subfragment used for a given reaction. The filter binding reaction was performed as described in Section 2. +, indicates that the DNA retained on the filter was 30–40-times the amount retained in the absence of β protein (background level). −, denotes the background level of retention, which is about 2% of total input. Abbreviations: A, *Ase*I; F, *Fsp*I; H, *Hin*dIII, and S, *Sty*I.

was not the case when the 593 bp *Hin*dIII–*Ase*I DNA subfragment was used (Fig. 3).

Digestion of the 815 bp *Ase*I–*Sty*I DNA segment with the *Fsp*I restriction enzyme renders DNA subfragments with 312 and 503 bp in length (Fig. 3). The β protein formed a complex with the 312 bp *Ase*I–*Fsp*I DNA subfragment, but not with the 503 bp fragment (Fig. 3). Therefore, we can infer that: (i) the β protein binds neither to the region encompassing gene β nor to the replication origin region, and (ii) the β protein does bind to a DNA segment that contains the promoter region of the orf α-gene β operon and the 5′-end of orf α (Fig. 4A) [2].

The β protein exhibits 92% identity to the OrfH protein of plasmid pAMβ1 [2,15]. Both β and OrfH proteins show homology with the Tn*3*/Hin family of sitespecific recombinase enzymes [2,6,7,15]. The biochemical properties of OrfH are unknown. Based on nucleotide sequence comparison with the DNA binding sites of the Tn*3*/Hin family of recombinases, Swinfield et al. [15] have proposed two putative DNA binding sites (namely pAMβ_1 and pAMβ_2) for protein OrfH (Fig. 4B). The region where the predicted pAMβ_2 (equivalent to coordinates 2,241 to 2,271 of plasmid pSM19035) was localized is 96% identical to the pSM19035 nucleotide sequence and corresponds to a subsegment of the 312 bp *Ase*I–*Fsp*I DNA fragment (Fig. 4A). On the other hand, the predicted pAMβ_1 DNA binding site (equivalent to coordinates 2,421 to 2,450 of pSM19035) is localized within the 503 bp *Fsp*I–*Sty*I DNA subfragment. We show that protein β was unable to bind to the 503 bp segment. This is not surprising because the region shows poor identity (56% homology) between pSM19035 and pAMβ_1 [2,15].

Computer based analysis of the 312 bp *Ase*I–*Fsp*I DNA fragment allows us to predict two putative binding sites for the β protein (I and II) that show some homology with each other and contain dyad axes of symmetry. Furthermore, the previous analysis (see above) can be extended to predict a promoter within these sites (see Fig. 4).

The nucleotide sequence where pAMβ_2 maps shows a few base pair differences with the predicted site I of pSM19035. Hence, it remains to be determined whether the β protein binds to the predicted sites I and II, and also whether it could bind to the pAMβ1 sequence.

Our finding that protein β is a dimer in solution is consistent with the presence of dyad symmetry in sites I and II, to which β should bind as a dimer. The promoter, predicted within the putative β protein binding sites I and II, would be responsible for the expression of orfs α and β. Therefore, protein β is likely to regulate its own synthesis by binding to sites overlapping with its own promoter, a common feature observed in the Tn*3*/Hin family of enzymes [6,7].

Acknowledgements This research was supported by grants-in-aid from the Spanish Ministry of Education and SBF 344/B5 to J.C.A., Bio93-131 from CICYT to F.R. and by an institutional grant from Fundación Ramón Areces to the Centro de Biología Molecular. We are indebted to W. Studier for providing plasmids and bacterial strains and to V. Kruft for performance of the protein sequencing. J.C.A. completed part of these studies while on a sabbatical leave at the

A

AseI

ATTAATTATTAGGGGGAGAAGGAGAGAGTAGCCCGAAAACTTTTAGTTGGCTTGGACTGAACGAAGTGAGGGAAAGGCTACTAAAACGTCGAGGGGCAGT

I

GAGAGCGAAGCGAACACTTGATCTTTTAAGTTGCTATCTTTTATA**GGTCAATAGAGTATACTTATTTGTCCTATTGATT**AGATAGCAGTATAAT**TAGCTTT**

pAMβ1₂

II

ATAGAGTAGGTCATTTAAGTTGAGCATAATAGGAGGATCAAGAATGAAAAAAATTTATTTATCGAGTTTTAGAAAATGACGAAGTGGTGGCTATTTTTAAT

FspI

GAGCAAGAATATGCGCA

B

	left arm of site I	central dinucleotide	right arm of site I
Tn3/Tn1000	G T C C G - A A A T A T T	**A T**	A A A T T A T - C G G A C
Gix (L + R)	T T A T C C A A A A - - C C	**T C**	G G - - T T T A C A G G A A
Tn917	G T A C C A T A A A - - C G	**A C**	C G - - T T T A T G G T A C
pIP404	G T A T C A A A A A - - T C	**C A**	C A - - T T T T T G A T A C
pAMβ1₁	T T A A C T A A A A - G T A	**G T**	G A A - T T T T T G A T T T
pAMβ1₂	A G G T C A T T A G - A G T	**A T**	A C T - T A T T T G T C C T
pSM19035	G G T C A A T A G A G T A T A C	**T T**	A T T T G T C C T A T T G A T T

Fig. 4. The β protein binding sequence and its comparison with the *res* site I consensus sequence of the Tn*3*/Hin family of recombinases (A) the nucleotide sequence of the 312 bp *Ase*I–*Fsp*I DNA segment is shown. The putative promoter is boxed. The ribosomal binding site (RBS) and the start codon of orf α are underlined. The predicted β protein binding sites (I and II) are in bold face letters The proposed pAMβ_2 binding site is denoted by a broken line. Converging arrows denote the regions of dyad symmetry (see text). (B) Sequence homology between site I recognition of DNA resolvases and left (L) and reight (R) arm of DNA invertases (reviewed in [6,7]) The nucleotide sequence of *res* site I and the sequence alignment have been reported elsewhere [2,7,15]. The promoter consensus regions (−10 and −35) are indicated.

laboratory of M. Salas at the Centro de Biología Molecular 'Severo Ochoa', Madrid, and wants to thank J.M. Lázaro for introduction to protein purification. The help and support of M. Salas is greately acknowledged.

REFERENCES

[1] Behnke, D., Tomich, P.K. and Clewell, D.B. (1980) Plasmid 4, 139–147.

[2] Ceglowski, P., Boitsov, A., Chai, S. and Alonso, J.C. (1993) Gene, in press.

[3] Ceglowski, P., Lurz, R. and Alonso, J.C. (1993) FEMS Lett. 109, 145–150.

[4] Brewer, B.J. (1988) Cell 53, 679–686.

[5] Futcher, A.B. (1986) J. Theor. Biol. 119, 197–204.

[6] Glasgow, A.C., Hughes, K.T. and Simon, M.I. (1989) in: Mobile DNA (Berg, D.E. and Howe, M.M. Eds.) pp. 637–660, American Society for Microbiology, Washington, DC

[7] Sherratt, D.J. (1989) in: Mobile DNA (Berg, D.E. and Howe, M.M. Eds.) pp 163–184, American Society for Microbiology, Washington, DC.

[8] Sambrook, J., Maniatis, T. and Fritsch, E.F. (1989) Molecular Cloning, a Laboratory Manual, second edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

[9] Laemmli, U.K. (1970) Nature 227, 680–685.

[10] Studier, F.W. and Moffatt, B.A. (1986) J. Mol. Biol. 189, 113–130.

[11] Krashnow, M.A. and Cozzarelli, N.R. (1983) Cell 32, 1313–1324.

[12] Alonso, J.C., Stiege, A.C., Dobrinski, B. and Lurz, R. (1993) J. Biol. Chem. 268, 1424–1429.

[13] Studier, F.W. (1991) J. Mol. Biol. 219, 37–44.

[14] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acid Res. 12, 387–395

[15] Swinfield, T.-J., Jannière, L., Ehrlich, S.D. and Minton, N.P. (1991) Plasmid 26, 209–221.